Supplementary Information

Supplementary Derivation of Plasmid Loss Rates

The age distribution of plasmid-free cells can be derived by dividing up the cell cycle into n small increments, and noting that the number x of plasmid containing cells at any increment follows

$$\frac{dx(1)}{dt} = c(2-p)x(n) - cx(1)$$
$$\frac{dx(j)}{dt} = cx(j-1) - cx(j) \qquad \text{for } 1 < j \le n$$

where c is a rate constant such that $n/c=T_x$, since each step takes on average $1/c$ time units, and it takes n step to complete a cell cycle of length T_x. We note that is the inherent generation time of the plasmid-containing cells, but that plasmid losses slows down the accumulated growth. This should in principle be accounted for experimentally, but since the effect is minute for virtually all plasmids, this is not practically important.

We further note that the total number of plasmid-containing cells is proportional to $(2\text{-p})^{t/Tx}$, since they multiply by a factor 2- p rather than double each generation. Thus the fraction g_i of cells at stage i follows from the product rule of differentiation

$$\frac{dg_i}{dt} = \frac{d\left[x(i)/x_{tot}\right]}{dt} = \frac{dx(i)}{dt} \times \frac{1}{x_{tot}} - \frac{\ln(2)(2-p)}{T_x}\frac{x(i)}{x_{tot}}$$

Combining the equations gives

$$\frac{dg_1}{dt} = \frac{n}{T_x}(2-p)g_n - g_1\left[\frac{n}{T_x} + \frac{\ln(2)(2-p)}{T_x}\right]$$
$$\frac{dg_j}{dt} = \frac{n}{T_x}g_{j-1} - g_j\left[\frac{n}{T_x} + \frac{\ln(2)(2-p)}{T_x}\right] \qquad \text{for } 0 < j \le n$$

which at the steady state produces

$$n(2-p)g_n = g_1\left[n + \ln(2)(2-p)\right]$$
$$ng_{j-1} = g_j\left[n + \ln(2)(2-p)\right] \qquad \text{for } 0 < j \le n$$

The second row simplifies to

$$g_i = \left(\frac{n}{n + \ln(2)(2-p)}\right)^{i-1} g_1$$

We next let the number of increments go to infinity, to recover continuous time, and note that $i/n = t/T_x$ where t now is the time in the cell cycle.

$$g(t) = \lim_{n\to\infty}\left(\frac{n}{n+\ln(2)(2-p)}\right)^{n\frac{t}{T_x}} g_1 = \left(e^{-\ln(2)(2-p)}\right)^{\frac{t}{T_x}} g_1 = (2-p)^{-t/T_x}\, g_1$$

We then calculate g_1 by noting that

$$\int_0^{T_x} g(u)du = g_1 \int_0^1 (2-p)^{-u/T_x}\, du = g_1 \left[\frac{e^{-\frac{\ln(2-p)}{T_x}u}}{-\frac{\ln(2-p)}{T_x}}\right]_0^{T_x} = g_1 \frac{(1-p)T_x}{(2-p)\ln(2-p)} = 1$$

Thus the solution is

$$g(t) = \frac{\ln(2-p)}{T_x(1-p)}(2-p)^{1-t/T_x}$$

We finally note that a shorter derivation is possible, using a heat equation approach and partial differential equations.

The changes in the population of plasmid-free cells in a time interval dt thus follow

$$x(t+dt) = x(t) + dt(1-p)\frac{\ln(2-p)}{T_x(1-p)}x(t)$$

and

$$\frac{dx}{dt} = \frac{\ln(2-p)}{T_x}x(t)$$

which, as a sanity check, directly results in the expected solution for $x(t)$ in the main text. This is already obvious from first principles (see above), and the advantage of the approach is that it also allows us to write a differential equation for the number of plasmid-free cells $y(t)$. Because only plasmid-containing cells that are at the end of the cell cycle can produce plasmid free cells, the accumulation of plasmid free cells $y(t)$ due to primary plasmid loss events and their subsequent growth can be described as follows

$$k = p\frac{\ln(2-p)}{T_x(1-p)}$$

$$\frac{1}{k}\frac{dy}{dt} = x(t) + \sum_{i=1}^{n} 2^{i-1}x(t-iT_y) \quad \text{for } t > nT_y$$

The first term corresponds to primary losses from plasmid-containing cells, and the terms in the summation correspond to the growth of plasmid-free cells generated by primary loss at previous times. The factor 2^{i-1} is the number of additional cells generated at time t for every cell created by primary losses at time $t - iT_y$. For example, if we divide the time axis into intervals of T_y, then in the first three intervals we have, respectively:

$$\frac{1}{k}\frac{dy}{dt} = x(t)$$
$$\frac{1}{k}\frac{dy}{dt} = x(t) + x(t - T_y)$$
$$\frac{1}{k}\frac{dy}{dt} = x(t) + x(t - T_y) + 2x(t - 2T_y)$$

Because $x(t)$ is known at all times, and simply decreases exponentially, solving for $y(t)$ is straightforward and becomes a series of piecewise differential equations, where the solution for one time interval becomes the initial condition for the next. The general and exact solution simply follows:

$$y(t) = \frac{px(0)}{1-p}\left(\left((2-p)^{t/\mathrm{Tx}} - 1\right) + \sum_{i=1}^{n} 2^{i-1}\left((2-p)^{\frac{t-i\mathrm{Ty}}{\mathrm{Tx}}} - 1\right)\right) \text{ for } n = \lfloor t/T_y \rfloor.$$

To account for initial populations of plasmid-free cells, we have to make assumptions about the distribution of cell cycle positions for those cells, i.e., the age structure of the population. Because those effects are small and we cannot reasonably postulate the exact mechanisms, we here assume those cells grow exponentially. An extra term $y(0)2^{t/T_y}$ can then be directly added to the solution above.

We can take an average slope by taking the derivative of $f(t)=x(t)/(x(t)+y(t))$, integrating this derivative over the interval between 0 and T_x, and then doing a first order Taylor expansion for small s

$$\frac{1}{T_x}\int_0^{Tx}\frac{df(t)}{dt}dt \approx -\frac{1}{2}f(0)(p + (1-f(0))(2-p)\Delta s\ln(2)) \approx -\frac{1}{2}f(0)(p + 2(1-f(0))\Delta s\ln(2))$$

Thus one can express the plasmid loss rate at short times by:

$$f(t) \approx f(0)(1 - p_{obs}t) \approx f(0)(1 - \frac{1}{2}(p + 2(1-f(0))\Delta s\ln(2))t$$

We note that because the plasmid-free cells created previously along the loss curve are similar to the plasmid free cells created before the beginning of the experiment (they are not identical because the age structure of the populations are slightly different) this expression can be used as an approximation throughout the plasmid loss assay using:

$$f(t) \approx f(t_0)(1-\frac{1}{2}(p+2(1-f(t_0))\Delta s \ln(2))(t-t_0)$$

As any part of the plasmid loss curve can be simply expressed by this expression, at least locally, the experimental challenge is to systematically minimize the contribution of the Δs term or to be able to determine that specific contribution to extremely high accuracy, in order to infer p.

Supplementary Protocol

A. Rapid measurement of plasmid loss rate

Materials:

- Coverslips
- Microscope slides
- Frame-seals (Bio-Rad, #SLF-0201 or SLF-0601)
- Low-melt agarose (MP Biomedical)
- LB broth
- Microscope with automated stage and 100x phase-contrast objective

Procedure:

1. Grow plasmid-containing strain in LB with selection overnight at 37°C.
2. Dilute culture by 1/20,000 into selective LB and incubate at 37°C with shaking until approximately OD_{600} of 0.3. A smaller OD_{600} value is ideal, but significantly lowers the density of microcolonies on the agarose pad during imaging.
    a. While culturing these cells, make agarose pads: attach one side of a Bio-Rad frame seal to a microscope slide and leave the other side with the protective film still on. Pipet molten 2% w/v low melt agarose in selective LB into the frame seal and place another microscope slide on top.
3. Centrifuge 1 ml of cells at 4,000 *g* and resuspend in 1 ml warm nonselective LB.
4. Immediately dilute tenfold into warm nonselective LB.
5. Carefully slide off the microscope slide, remove excess agarose by peeling off the protective film, and slightly trim the edges of the agarose pad with a scalpel to prevent later bowing of the coverslip. Do this step very quickly; it should take no longer than 30 seconds.
6. Pipet 5 microliters of the warm diluted culture onto the agarose pad. Tilt the microscope slide so that the droplet evenly wets the agarose pad. When the pad dries, immediately seal with a coverslip.
7. Incubate the diluted culture at 37°C.
8. Repeat steps 5-6 for replicates and time intervals; to ensure that the culture does not cool significantly, pipet from the culture at the incubator and return immediately to 37°C.
9. For each time point, incubate microscope slides for 30 minutes at 37°C. Afterward, remove from 37°C and place at room temperature for another 30 minutes where after they can be stored at 4°C before imaging. Residual growth still occurs at room temperature as the microscope slide cools down, but immediately cooling the microscope slide to 4°C causes movement of the agarose pad, excessive condensation, and poor imaging.

10. To image microscope slides, place agarose pad over an oil-immersion phase contrast objective at 100x magnification. Using an automated stage, acquire 100 frames in a 10x10 non-overlapping grid pattern.
11. Repeat for all replicates and time points.
12. Tally all microcolonies manually for the presence of plasmid-free cells.

General remarks: The number of microcolonies to measure is variable depending on the plasmid loss rate. In general, one should take as many images as possible but this is limited by a variety of practical limitations. As a rule of thumb, for measuring rare populations the sample size for each observation should be higher than $3/y$ where y is the fraction of plasmid-free cells. This is derived from solving the probability of observing at least one plasmid loss event ($M > 0$) in a sample size of *N*:

$$p(M = m) = \frac{(Ny)^m \exp(-Ny)}{m!}$$

$$p(M > 0) = 1 - p(0) = 1 - \exp(-Ny)$$

For any given statistical power $1 - p(0)$, *N* can be found to be $N > -\ln(p(0))/y$. For $1 - p(0) = 0.95$, $N > 3/y$. Therefore, for $y = 0.005, N > 600$. In our experiments, we accordingly measure between 1000-2000 microcolonies per replicate.

B. Fluctuation-based measurement of plasmid loss rate

Materials:

- Multichannel pipette + pipetting reservoirs
- Round-bottom 96 well plates
- Breathe-Easy plate sealing membrane (Sigma-Aldrich)

Procedure:

1. Grow strain overnight at 37°C with shaking in selective LB broth
2. Dilute cells by at least 10^8 into a final large volume of LB broth. Approximately 110 ml is needed for pipetting into a 96 well plate.
3. Pour diluted cells into reservoir and load 100ul into each well with a multichannel pipette.
4. Seal with Breath-EZ film. Be sure to take off the top layer.
5. Repeat from step 2 for desired number of replicates.
6. Make an extra plate for population estimates. Alternatively, some wells can be individually sampled; these are thrown out of the analysis.
7. Wait 2.5 to 3 hours for mini-R1 containing plasmids. For other plasmids that are more stable the time must be increased. The goal is to have only a subset of the wells

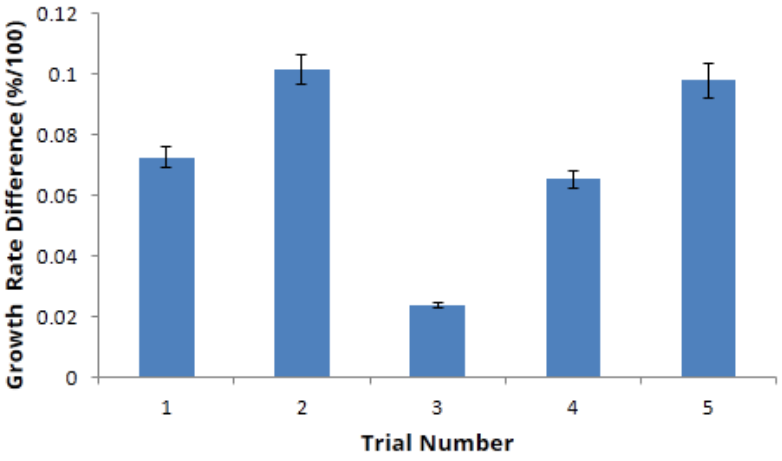
containing plasmid-free cells in order to exploit Poisson statistics. Hence, the waiting time is not particularly critical, but it is important to not let the experiment continue for too long, as all the wells will be populated with plasmid-free cells.

8. Remove the plates from the incubator and carefully peel off the sealing membrane
9. Make dilutions from each well of the extra plate and plate on LB agar to estimate population size. Alternatively, some wells from all plates can be individually sampled – these wells are then subsequently taken out of the analysis due to possible removal of plasmid-free cells. Culture at 37°C. In our experience, most of the experimental error comes from the estimation of population size.
10. Make a reservoir of chloramphenicol at 3x selective concentration in LB broth – this is the counterselecting LB broth. Using chloramphenicol from ethanol stocks is hard to pipette accurately and difficult to mix.
11. With a multichannel pipette, draw up 50 ul of the counterselecting LB broth. Hover over a row of wells and carefully dispense, making sure to avoid splashes. Angling the multichannel pipette such that the broth dispenses against the wall of the well reduces splashing. Alternatively, you can immerse the pipette tips into the wells and change tips at every row but this is somewhat wasteful.
12. Seal all plates with new Breathe-Easy membranes and incubate overnight at 37°C. Because the cells are starting from a very small initial population size, it will take slightly longer than overnight to reach saturation (30 generations is approximately 10^9).
13. The next day, count the fraction of clear wells (p_0) – these are the ones that did not have any plasmid-free cells arise.
14. Count colonies on plates to estimate population size N. Calculate loss rate as $p_{loss} = -\log(p_0)/N$.

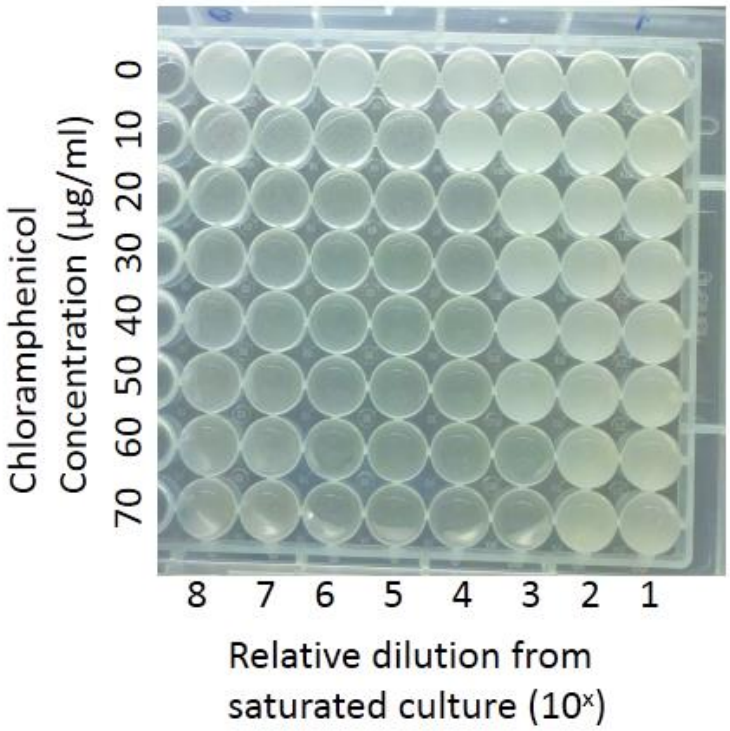
General remarks: Several optional controls can be made:

A. *Initial dilution control:* In step 2, one extra full plate can be immediately cultured overnight with no counterselection in order to assess the rate of negative cell loading. In our experience, dilutions of up to $5 * 10^9$ still resulted in all wells being positively loaded. Plain LB broth can also be added to one plate in step 11 for the same purpose.
B. *Counterselecting antibiotic concentration:* Add varying amounts of antibiotic to a diluted culture from step 1 to determine appropriate antibiotic concentration. In our experience the strong $P_{LlacO-1}$ promoter is sufficiently strong to allow growth even up to chloramphenicol concentrations of 60 ug/ml.

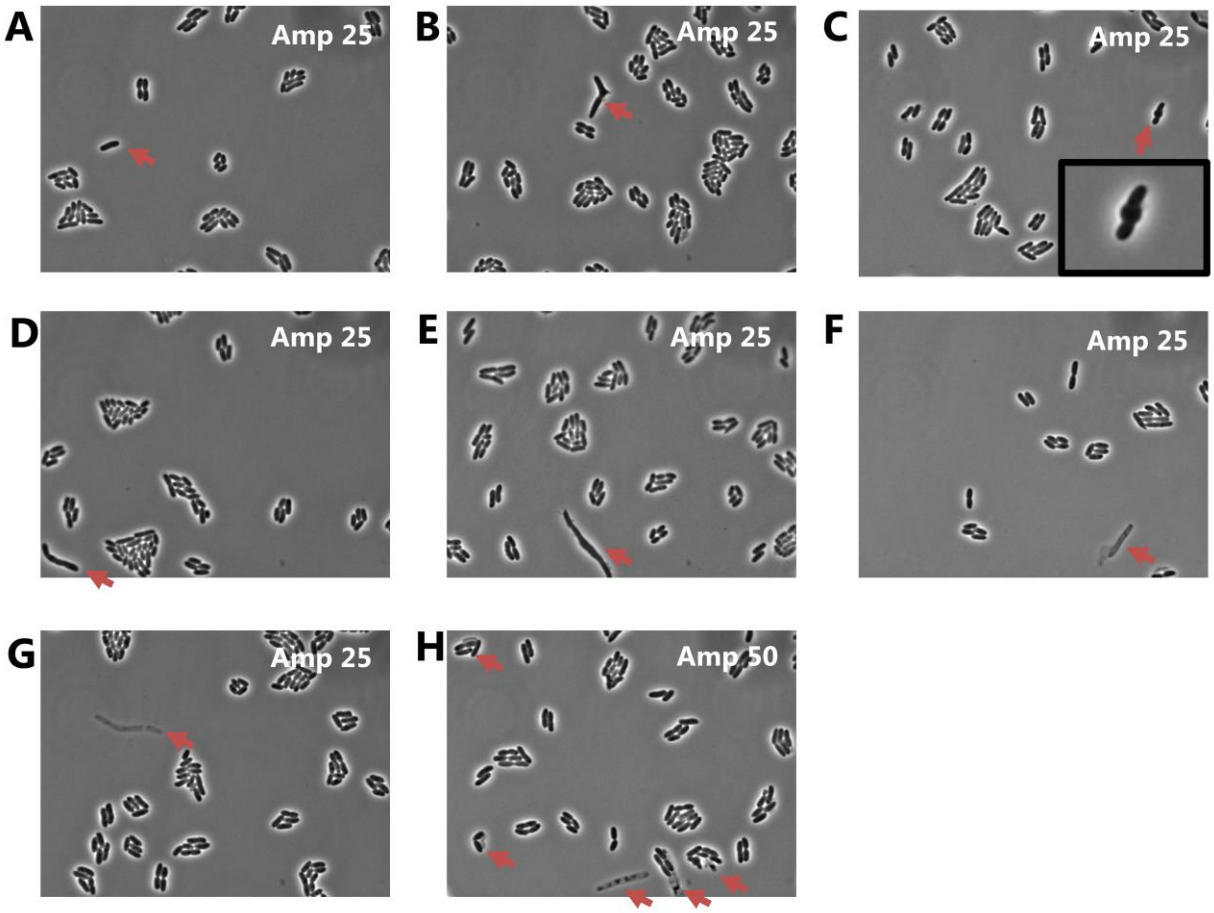
C. *Calculating error bars:* Standard deviations can be calculated by using standard error propagation techniques: $\frac{\sigma_{ploss}^2}{<p_{loss}>^2} = \frac{\left(\frac{d\left(-\frac{\ln(p_0)}{N}\right)}{dN}\right)^2 \sigma_N^2 + \left(\frac{d\left(-\frac{\ln(p_0)}{N}\right)}{dp}\right)^2 \sigma_{p_0}^2}{\left(-\frac{\ln(p_0)}{N}\right)^2} = \frac{\sigma_N^2}{<N>^2} + \frac{\sigma_{p_0}^2}{(p_0 \ln(<p_0>)^2}$.



Supporting Figure S1. Measurement of plasmid burden of pBTCL89. OD_{600} measurements of mini-R1 plasmid-containing and plasmid-free strains were carried out in an incubated Victor3V plate reader at 37°C (N=48 for each trial). 200 µl of a 1000-fold dilution from saturated cultures were dispensed in each well. A fit against the equation $\exp(bt - a) + c$ was made for each well, whereafter b was averaged and then a corresponding doubling time was calculated by $\tau = \ln(2)/b$. c is the OD_{600} offset of the plate reader as determined by the beginning time points.



Supporting Figure S2. Calibration of antibiotic concentration used for counterselection. A saturated culture of the mini-R1 plasmid pBTCL89 was diluted in LB medium at a volume of 100 µl per well. 50 µl of LB medium at containing various amounts of chloramphenicol was added to each well for final concentrations as indicated. The plate was then grown overnight at 37°C on a plate shaker. The plasmid-free background strain can grow in 60 µg/ml chloramphenicol-containing LB medium from high dilution (10^6-10^8) to saturation (not shown).



Supporting Figure S3. Sample micrographs of microscopy-based plasmid loss screening. (A-G) Various single-cell morphologies of plasmid-free cells affected by ampicillin at 25 µg/ml indicated by the red arrows. Other microcolonies are otherwise appear healthy and therefore plasmid-containing. (H) At 50 µg/ml, many single cells are affected but many microcolonies also contain some affected cells, indicating that the antibiotic selection strength was too high. An inset in (C) is provided.

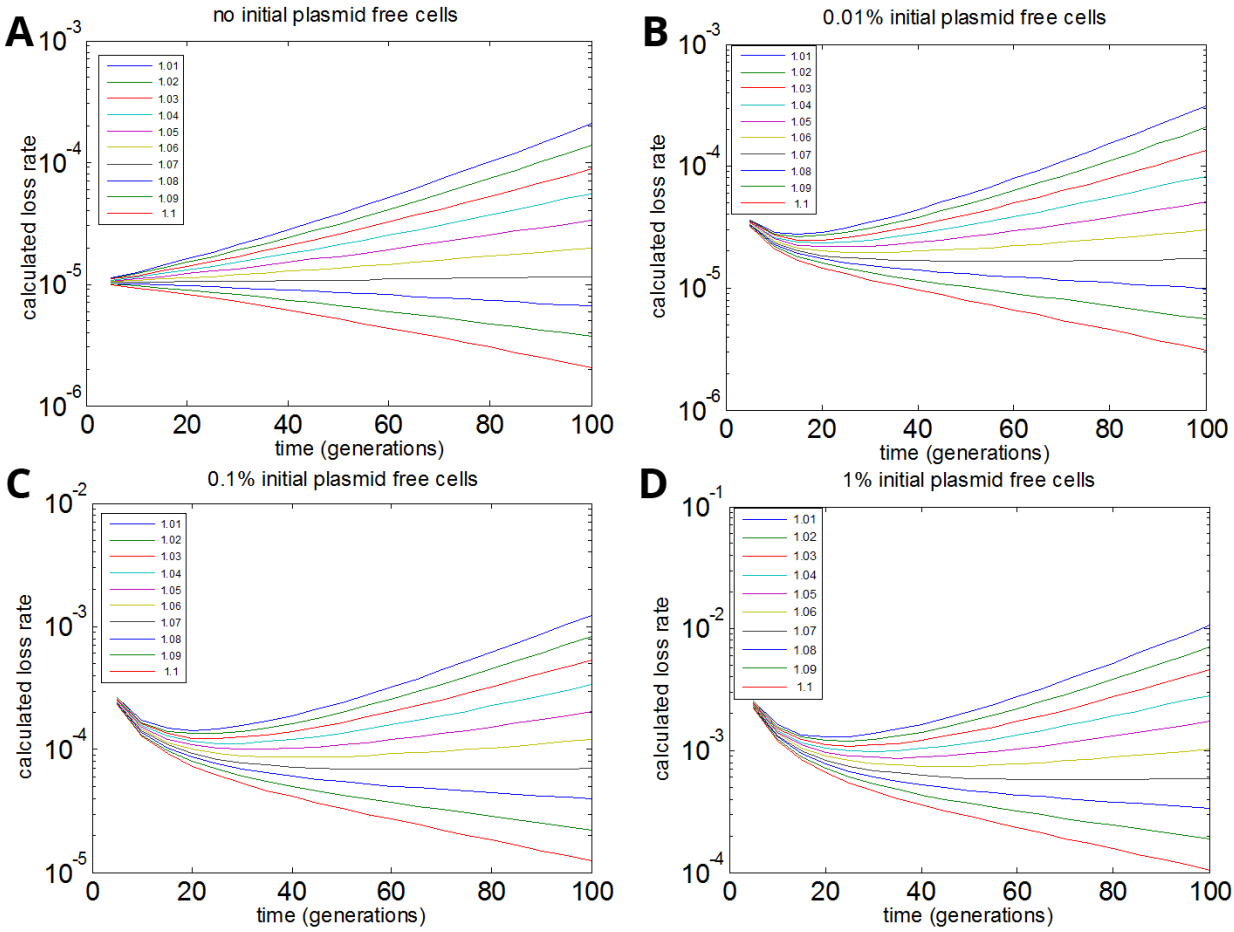


Figure S4. Variations in the estimated plasmid loss rate due to length of experiment, differences in growth burden estimates, and initial plasmid-free fraction. Mock data is generated by solving the plasmid loss differential equation in the main text (p_{loss} = 0.00001 and k_y/k_x = 1.05) and generating values for the fraction of plasmid-containing cells to 100 generations. The initial fraction of plasmid-free cells are (A) 0%, (B) 0.01%, (C) 0.1%, and (D) 1%. The plasmid loss rate is calculated with "method 2" as described in Boe and Rasmussen, *Plasmid* 1996. Each curve represents a different estimate of growth burden used for the calculation of plasmid loss rate.

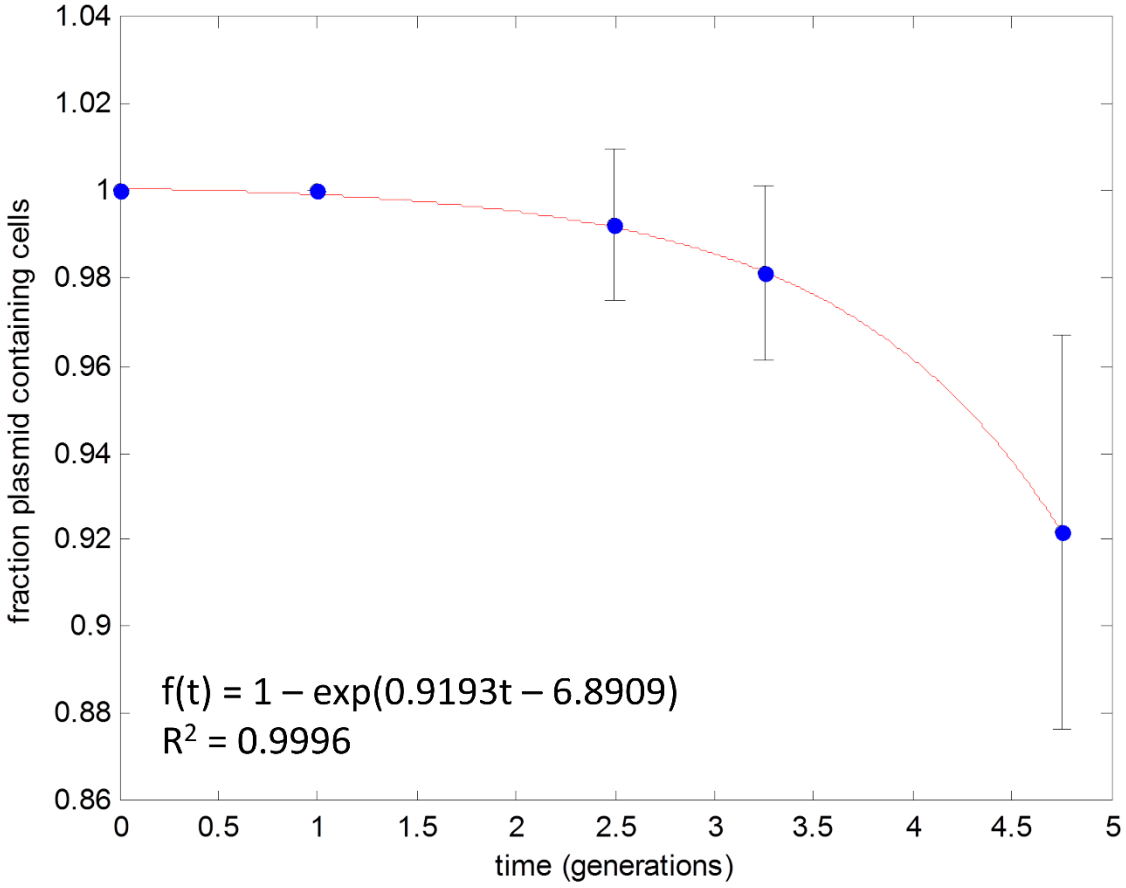


Figure S5. Measuring plasmid loss past one generation. A culture containing pBTCL89 was screened for plasmid loss by standard plating assays. The time-course fraction of plasmid-containing cells was fitted to an exponential function f(t) = 1 - exp(at + b).

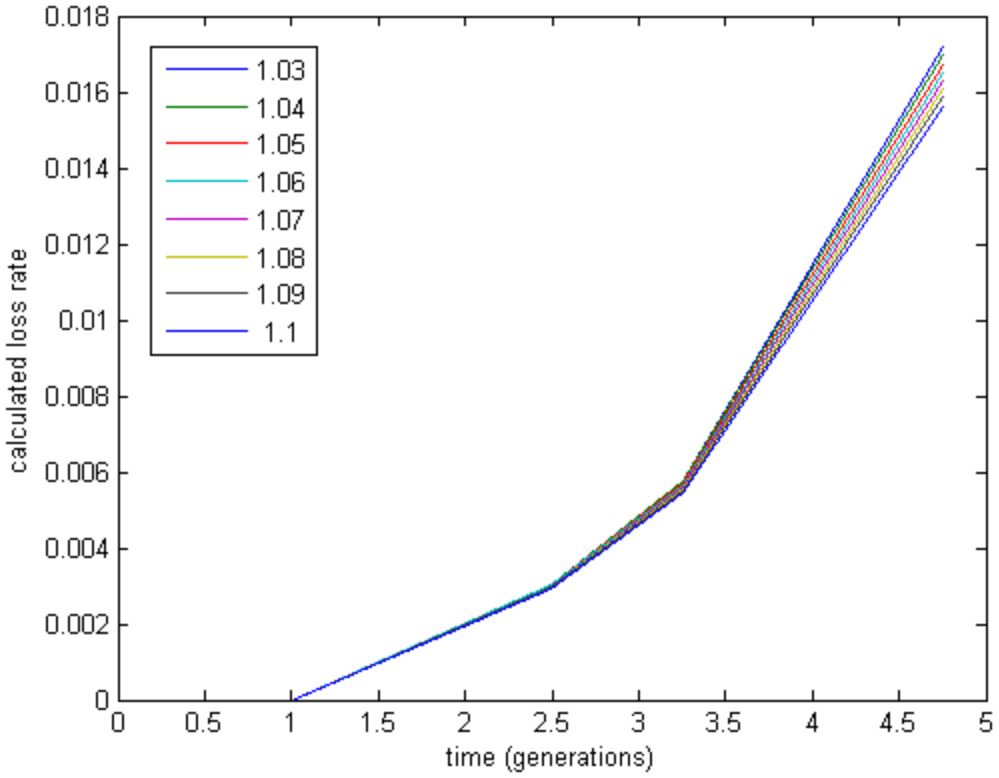


Figure S6. Variation in estimated plasmid loss rate past one generation. Plasmid loss rates were calculated from data in Figure S5, accounting for growth burdens using "method 2" as suggested in Boe and Rasmussen, Plasmid 1996. Each curve represents a different estimate of growth burden used for the calculation of plasmid loss rate.

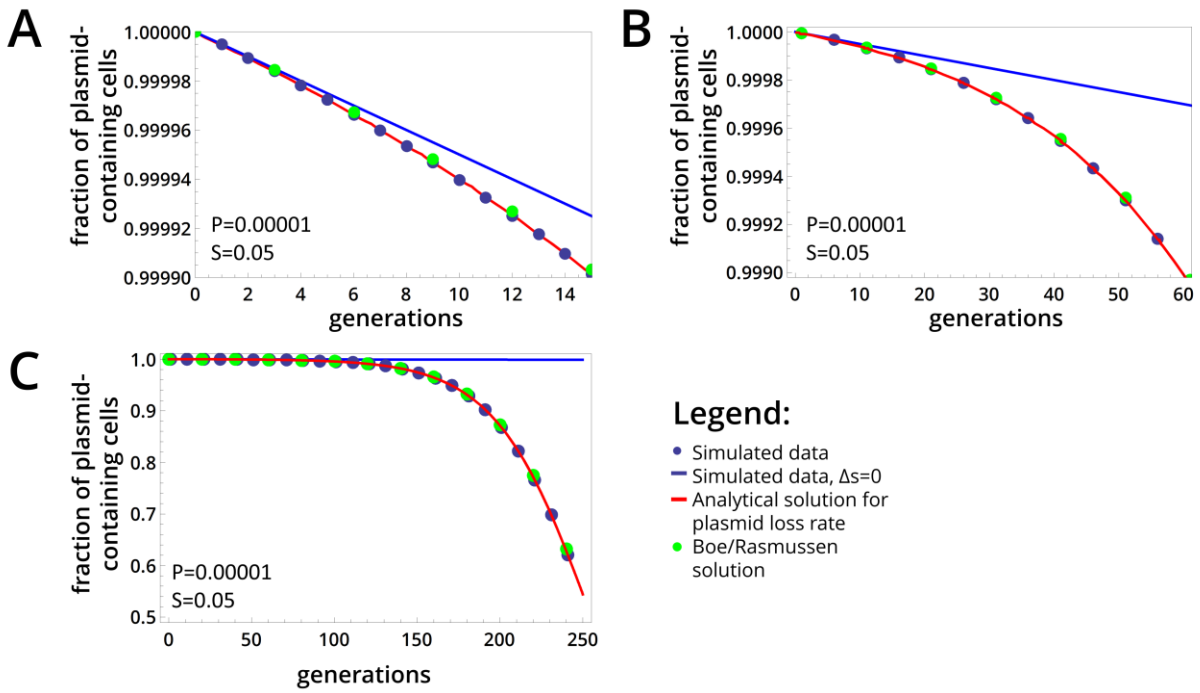


Figure S7. Equivalence between plasmid loss models. The fraction of plasmid-containing cells with dynamics is plotted as described in Boe and Rasmussen, *Plasmid* 1996 by solving Equation 7 in their "method 2" section for the fraction of plasmid-containing cells. The analytical solution for the plasmid loss rate dynamics as described in the main text is also plotted, showing excellent agreement with simulated data. To gauge relative contributions from primary plasmid loss events and its subsequent plasmid-free differential growth, simulated data corresponding to $\Delta s=0$ is also plotted. Different time scales are represented: (A) 15 generations, (B) 60 generations, and (C) 250 generations. Parameters: $p=0.00001$, $\Delta s=0.05$, starting with a pure plasmid-containing population.

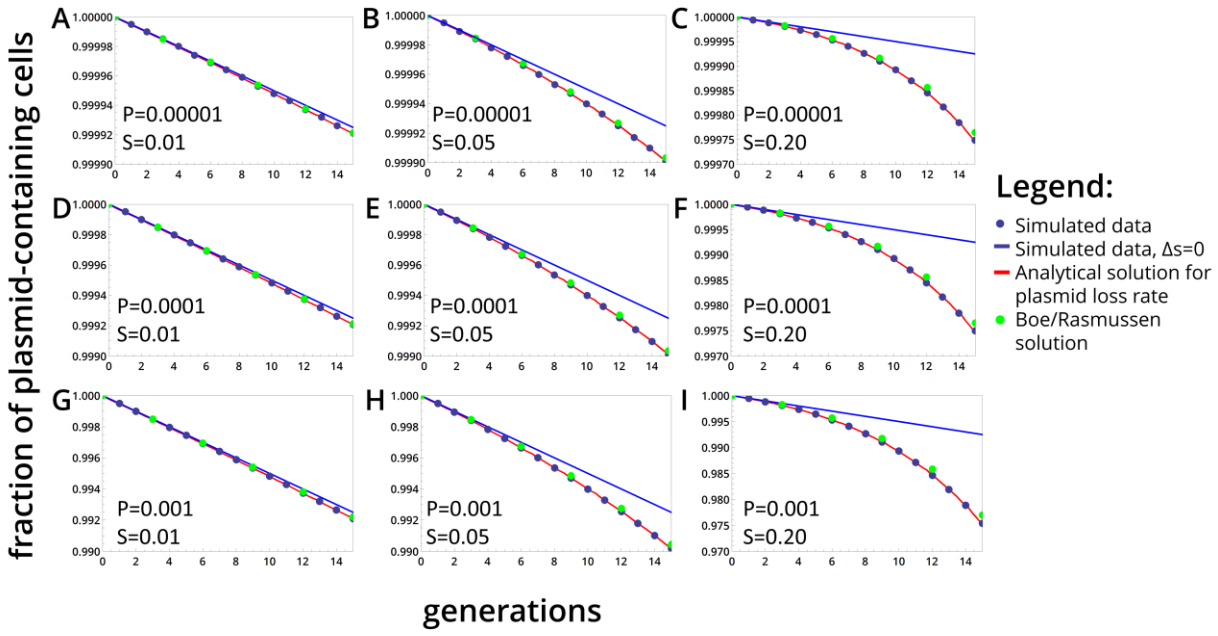


Figure S8. Equivalence between plasmid loss models for different parameter regimes. The fraction of plasmid-containing cells with dynamics is plotted as described in Boe and Rasmussen, *Plasmid* 1996 by solving Equation 7 in their "method 2" section for the fraction of plasmid-containing cells. The analytical solution for the plasmid loss rate dynamics as described in the main text is also plotted, showing excellent agreement with simulated data. To gauge relative contributions from primary plasmid loss events and its subsequent plasmid-free differential growth, simulated data corresponding to $\Delta s=0$ is also plotted. A range of parameters are considered for 15 generations of growth: (A) $p=0.00001$, $\Delta s=0.01$, (B) $p=0.00001$, $\Delta s=0.05$, (C) $p=0.00001$, $\Delta s=0.20$, (D) $p=0.0001$, $\Delta s=0.05$, (E) $p=0.0001$, $\Delta s=0.05$, (F) $p=0.0001$, $\Delta s=0.20$, (G) $p=0.001$, $\Delta s=0.01$, (H) $p=0.001$, $\Delta s=0.05$, and (I) $p=0.001$, $\Delta s=0.20$.